Comparative study on photometric normalization algorithms for an innovative, robust and real-time eye gaze tracker

J Real-Time Image Proc (2013) 8:21–33
DOI 10.1007/s11554-011-0217-6

SPECIAL ISSUE

Comparitive study on photometric normalization algorithms for an innovative, robust and real-time eye gaze tracker

Antonino Armato · Antonio Lanatà · Enzo Pasquale Scilingo

Received: 31 January 2011 / Accepted: 16 July 2011 / Published online: 9 August 2011

Abstract Eye gaze trackers (EGTs) are generally developed for scientific exploration in controlled environments or laboratories and data have been used in ophthalmology, neurology, psychology, and related areas to study oculomotor characteristics and abnormalities, and their relation to cognition and mental states. The illumination is one of the most restrictive limitation of the EGTs, due to a problem of pupil center estimation during illumination changes. Most of the current systems, indeed, work under controlled illumination conditions either in dark or indoor environments, e.g. using infrared sources or conforming the sources of light to fixed levels or pointing directions. This work is focused on exploring and comparing several photometric normalization techniques to improve EGT systems during light changes. In particular, a new wearable and wireless eye tracking system (HATCAM) is used for testing the different techniques in terms of real-time capability, eye tracking and pupil area detection. Embedding real-time image enhancement into the HATCAM can make it an innovative and robust system for eye tracking in different lighting conditions, i.e. darkness, sunlight, indoor and outdoor environments.

A. Armato
Department of Information Engineering,
University of Pisa, via Caruso, 16, 56122 Pisa, Italy

A. Lanatà · E. P. Scilingo (✉)
Interdepartmental Research Center "E. Piaggio"
and Department of Information Engineering,
University of Pisa, via Diotisalvi, 2, 56126 Pisa, Italy
e-mail: e.scilingo@centropiaggio.unipi.it

A. Lanatà
e-mail: a.lanata@centropiaggio.unipi.it

1 Introduction

Eye gaze trackers (EGTs) have been and are currently used in many research areas, such as marketing and advertising, as well as in human factors engineering to evaluate computer interfaces and web sites, Duchowski [9]. Moreover, EGTs have been studied as input devices for computer interfaces especially for people with disabilities. A specific field of application, on which we are currently focusing our efforts, is the study of oculomotor characteristics and abnormalities, and their relation to cognition and mental states. EGTs rely on an analytical relationship between relative eye movements and the point of gaze over time, and all the current systems which can be found in literature are differentiated by the method which are used to calculate eye movements. There are several eye-movement measurement methods based on mechanical or optical referenced objects mounted either on scleral contact-lens or directly on the eye (e.g. search coil), which result very precise although invasive, Duchowski [9]. Other methods, based on electrooculogram (EOG), Robinson [33], Young and Sheena [42], are affected by several disadvantages such as obtrusiveness, low resolution, drift, noise, electromyogram (EMG) artifacts, in addition to discomfort for the patient who is limited in his normal activities, DiScenna et al. [6]. A specific class of EGTs, that tries to overcome these drawbacks, is based on video-oculography (VOG), which is a method for tracking eye movements by digitally processing video images of the eyes. Pupil position and iris landmarks are detected by means of image processing algorithms and used to calculate eye rotation angles and its center as well. Currently, EGTs can be split into two different classes, remote and wearable. Remote eye-tracking systems require the user to keep their head still, thus making the systems unsatisfactory for prolonged

use in interactive applications, even though they achieve an optimal accuracy, specially for clinical usage, Meyer et al. [27]. The second class, commonly referred to as head-mounted EGTs, uses eye tracking systems, i.e. cameras or other equipments, mounted directly on the user head. The development of these wearable EGTs has opened new scenarios where the user is free to move overcoming the spatial limitations imposed by the remote systems, Zhu and Ji [43]. In addition, their use outside controlled environments allows investigating eye-movement measurements during natural tasks, e.g. driving cars, Land and Lee [21], hand washing, Pelz et al. [30], walking through a room cluttered with obstacles, etc.

On the other hand, the extended workspace of EGTs introduces some difficulties in estimating the pupil center during illumination changes and consequently to the detriment of robustness. At the present time, the eye localization, independently of illumination, is still a crucial issue. Current systems, indeed, are constrained to be used in controlled environments where illumination is kept constant. However, in the last few years many efforts have been provided to overcome this limitation. One of the most significant approach is based on detecting eye tracking using active infrared (IR) illuminators. In particular, spectral (reflective) properties of the pupil, under near-IR illumination, are exploited to maximize the image-contrast between the pupil and the background. Based on this principle, several tracking techniques, Morimoto et al. [29] and Morimoto and Mimica [28], have been developed as well as some commercial eye trackers have been produced, LcTechnologies [22] and Anon [1].

For example, OpenEyes, developed by Li et al. [23], which is a low cost head-eye tracker system composed of two CCD cameras mounted on a pair of safety glasses. In this system, one camera captures an image of the eye while the other captures an image of the scene. An IR LED is placed in axis with the eye-camera producing an illumination which allows, together with an ad-hoc algorithm based on Random Sample Consensus (RANSAC) paradigm, the discrimination of the pupil from the rest of the eye, Dongheng et al. [7]. Also Babcock et al. [2] introduced other EGTs based on an IR mirror, an IR camera and an IR illuminator while Franchak et al. [12] introduced very lightweight systems suitable for children based on two light cameras and an IR LED to produce a reference point on the pupil.

However, even though uniform and controlled IR illumination eliminates uncontrolled specular reflection and it is not perceivable by the user, it does not permit the natural aversion response which protects eyes against retinal injuries when viewing very bright light sources, San Agustin et al. [34]. Even though low-power IR LEDs are, generally, employed to avoid injuries, the irradiance level must be kept less than 10 mW/cm^2 for not chronic IR exposure in the range of 720–1,400 nm, Sliney et al. [36]. Moreover, the use of IR illuminator systems, especially for high-sensitive subjects (e.g. children), can produce reddening and lachrymation. Furthermore, in subjects with eye glasses the lens disturb the IR light thus showing very weak pupils. Finally, a large variation of bright light sources can produce a diminished image of the pupil or even its disappearance. These limitations in eye-tracking methods using IR illuminators impose stable lighting conditions, and therefore a restriction of the fields of application. Therefore, it is very important to develop an EGT able to robustly and accurately track eyes under several illumination conditions and during their changes.

This work is focused on a comparative performance evaluation of the most relevant and widely used photometric normalization techniques, through a wearable EGT system. More specifically, eye gaze accuracy, pupillometry, and algorithm execution time have been evaluated in three different illumination conditions: darkness, laboratory and sunlight conditions.

2 HATCAM—wearable EGT system

In this work, we use a new wearable and wireless eye tracking system (HATCAM) which can be tailored to both adults and children. It is comprised of only one lightweight camera which was able to capture, by means of a mirror, the eyes of the subject and the scene in front of him, simultaneously. The system can be used indoor and outdoor environments. It exhibits the following proprieties:

1. wearability,
2. minimal obtrusiveness,
3. eye tracking and pupillometry capabilities,
4. lightweight below 100 g,
5. wireless communication.

The system configuration is shown in Fig. 1.

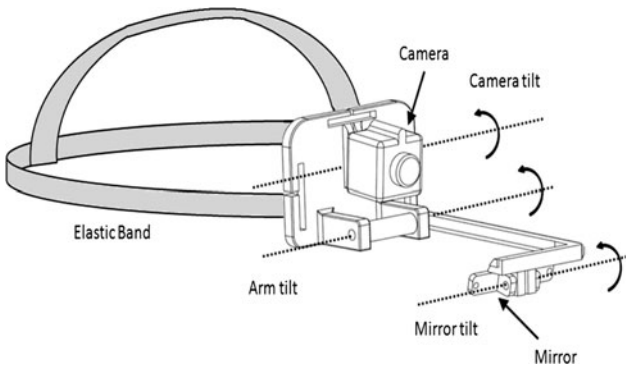


Fig. 1 HATCAM configuration

In detail, the system is comprised of a wireless CMOS camera (CP294) having low weight (20 g), low size (2 × 2 × 2 cm), and an A/V transmitter up to 30 m of distance. The camera has a resolution of 628 × 586 pixels with *F*2.0, *D*45° optic, and an acquisition frequency of 25 frames per second (fps). The original lens of the camera was removed and substituted with a wide-angle lens without IR filter. This operation allows enlarging the view angle and acquiring IR components, which emphasizes the contrast between pupil and iris. This system is able to capture simultaneously, without latency, the visual scene in front of the subject and the position of the eyes. This is achieved using a mirror (4 × 0.6 cm) placed on a shaft linked to the head (see Fig. 1). Tilt and shaft of the mirror and the camera orientation can be tailored to the forehead profile of the user (see Fig. 1).

3 Eye tracking method

This section deals with the processing techniques used to detect the center of the eye and how its movements are mapped into the image plane. VOG method involves visible spectrum imaging. This technique is a passive approach that captures ambient light reflected from the eye. The mounted camera is modified to acquire also the near IR components of natural light. Therefore, the system keeps the advantages of IR illumination in increasing the contrast between pupil and iris, and at same time preventing any possible injuries due to artificial IR illuminators. Indeed, as the bandwidth includes both visible light and IR components, no illuminators are needed. Figure 2 shows the block diagram of the image processing. In this diagram, image of the eyes together with the scene are inputs of the processing chain. More in detail, the first processing block implements the eye extraction algorithm, the second applies the photometric normalization algorithm of illumination, the third block extracts the pupil contour and implements the fitting algorithm. At this point, the center of the eye is detected. The last block is constituted of an appropriate function, which maps the eye center and movements into the image plane. This function is generally named "*mapping function*". In the next sections, each block is described more in depth.

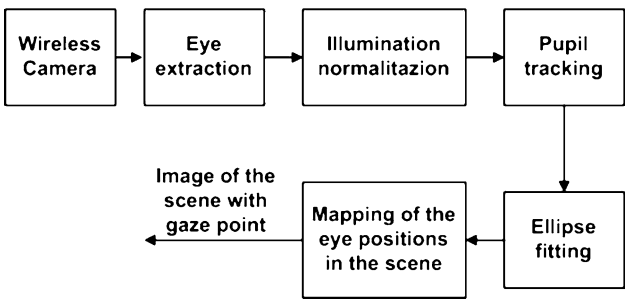


Fig. 2 Block diagram showing all the algorithmic stages of the processing of eyes and outside scene

3.1 Eye extraction

Figure 3 shows how the HATCAM is able to acquire simultaneously the eyes of the user and the scene in front of him using the mirror. Eye extraction procedure is constituted of visual inspection of the first video frame, in which a rectangular area including the eye is manually selected, in order to delimit the smallest region including the eye (see Figs. 3 and 4). This region is called region of interest (ROI). Since the eye tracking is a head-mounted system, the ROI does not change thourought the experiment. In addition, only the red-image component is converted in gray scale and used as input to the other processing blocks (see Fig. 4), as this component is specifically helpful in enhancing the contrast between pupil and background.

3.2 Photometric normalization techniques

Seven photometric normalization (photonormalization) techniques were applied and compared. The purpose of the illumination normalization is to reduce or eliminate some

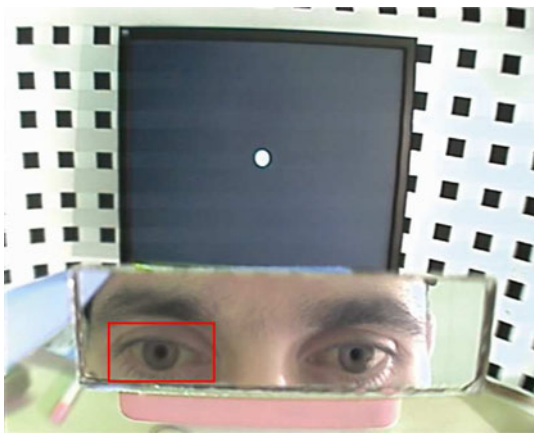


Fig. 3 Example of single frame captured by the camera. The rectangular area marked by *red* represents the ROI

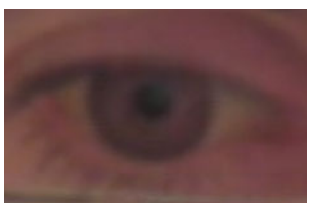


Fig. 4 Red component of the ROI

variations in the captured eyes due to different conditions of illumination. It normalizes and enhances the eye image to improve the recognition performance of the system. The photometric normalization consists in removing the mean of the geometrically normalized image and scaling the pixel values by their standard deviation, estimated over the whole cropped image, Ranawade [32]. All the photometric normalization techniques applied in this work are based either on homomorphic filtering or histogram equalization. In recent years, several approaches have been proposed to solve the brightness variation issue in different fields of application such as face recognition, Chen et al. [5], or automotive scenarios, Marsi and Saponara [26].

Six out of the seven explored photonormalization techniques are based on the Retinex theory (from the words "*retina*" and "*cortex*", suggesting that both eye and brain are involved in the processing) developed by Land and McCann [20]. This theory is based on color constancy assumption which ensures that the perceived color of objects remains relatively constant under varying illumination conditions. Land and his colleagues assume that the stimulus is not the result of the light source and surface reflectivity only, but that the visual system processes the stimulus, integrating the spectral radiance and generating a ratio of integrated radiance of any region of the scene with that of the brightest region. This stimulus is called *lightness*. This model eliminates the effect of a non uniform illumination and is completely independent of any a priori knowledge of the surfaces reflectance and light source composition.

Two major assumptions underly this theory:

- The human visual system performs the same computation independently of the color channel,
- In each channel, the signal intensity is proportional to the product of the illumination source and the surface reflectance, which is determined by the object characteristics.

According to this theory, the image intensity $I(x, y)$ can be simplified and formulated as follows:

$$I(x,y) = R(x,y)L(x,y) \tag{1}$$

where $R(x, y)$ is the reflectance and $L(x, y)$ is the luminance at each point (x, y). The luminance L is assumed to contain low frequency components of the image while the reflectance R mainly includes the high frequency components of the image. This assumption can be easily understood if considering that R generally varies much faster than L does, in most parts of the image with a few exceptions, e.g. shadow boundaries, where L changes remarkably. In addition, in a real world scene, the illumination can dynamically change much more than the reflectance. In conclusion, there is a widely accepted statement about human vision, which assumes that human eyes respond to local changes in contrast to global brightness levels.

Below a detailed description of the seven photonormalization techniques, used in this work, is reported.

3.2.1 Single scale retinex (SSR)

The latest version of Land Retinex theory was implemented by Jobson et al. [16] as SSR. The main idea of the algorithm is to process the image through a class of center surround functions where each output value of the function is determined by the corresponding input value (center) and its neighborhood (surround). The center is defined as each pixel value and the surround is a Gaussian function. The mathematical form of the SSR is given by:

$$\mathrm{SSR}(x,y) = \log(I(x,y)) - \log[F(x,y) \times I(x,y)] \tag{2}$$

where $\mathrm{SSR}(x, y)$ is the Retinex output, $I(x, y)$ is the input image, and $[F(x, y) \times I(x, y)]$ is a convolution product between $I(x, y)$ and $F(x, y)$. This latter function is a simple linear filter with Gaussian kernel:

$$F(x,y) = k\mathrm{e}^{\frac{-r^2}{\sigma^2}} \tag{3}$$

where $r = \sqrt{x^2 + y^2}, \sigma$ is the standard deviation of the filter, empirically determined, and controls the amount of spatial detail that is retained in terms of pixels, and k is a normalization factor that keeps the area under the Gaussian curve equal to 1.

3.2.2 Multi scale retinex (MSR)

Although the SSR algorithm produces good results with a properly selected Gaussian filter, it is still limited by an important weakness: at large illumination discontinuities caused by strong shadows that are casted over the scene halo effects are often visible in the computed reflectance, Jobson et al. [17]. To avoid this inconvenient, Jobson extended the SSR algorithm to a multi scale form, where Gaussian filters with different widths are used and the output combines different implementations of distributed SSR algorithms to compute the final illumination invariant image representation. In detail, the MSR output is simply a weighted sum of the outputs of several different SSR outputs. Mathematically,

$$R^i_{\mathrm{MSR}} = \sum_{n=1}^{N} w_n R^i_n \tag{4}$$

where N is the number of scales, R^i_n is the ith component of the nth scale, R^i_{MSR} is the ith spectral component of the MSR output, w_n and is the weight associated with the nth

scale. The only difference between $R(x, y)$ and $R_n(x, y)$ is that the surround function is given by

$$F_n(x, y) = ke^{\frac{-r^2}{\sigma_n^2}} \tag{5}$$

3.2.3 Single scale selfquotient image (SQI)

The SQI algorithm is an extension of the quotient image (QI), Wang et al. [40, 41]. QI is a concept introduced by Shashua and Riklin-Raviv [35]. It is based on the Lambertian model where the image can be described by the product of the albedo and the cosine angle between a point light source and the surface normal:

$$I(x, y) = \rho(x, y)\mathbf{n}(x, y)^{\mathrm{T}} \cdot \mathbf{s} \tag{6}$$

where ρ (x, y) is the albedo, $\mathbf{n}(x, y)^{\mathrm{T}}$ is the surface normal (shape) of the object (same for all objects of the class), and $\mathbf{s}$ is the light source direction, which may vary arbitrarily. QI of two objects belonging to the same class is defined as the ratio between the albedo of each object. The albedo represents the diffuse reflectivity or reflecting power of a surface. It is defined as the ratio of reflected radiation from the surface to incident radiation upon it. Being a dimensionless fraction, it may also be expressed as a percentage, and is measured on a scale from zero for no reflecting power of a perfectly black surface, to 1 for perfect reflection of a white surface.

The SQI algorithm is defined as follows:

$$Q(x, y) = \frac{I(x, y)}{F \times I(x, y)} \tag{7}$$

where $Q(x, y)$ is the result of the algorithm, $I(x, y)$ is the input image as defined in QI and F is a smoothing filter. Unlike the QI, the SQI has several advantages, among which the most relevant for our purposes is that no image training processes are required which implies that it can be used as a pure pre-processing algorithm. A crucial issue to be carefully addressed in this technique is the filter size. If the scale is too small, indeed, the filtered image will be close to the original image, but if the filter is too large, the filtered image will be more or less constant, and the input image will be normalized with its mean value.

3.2.4 Multi scale selfquotient image (MSQ)

The MSQ technique exhibits similarities to MSR, but unlike it, this technique uses an anisotropic filter for the smoothing operation. Moreover, MSQ allows overcoming the limitation of the filter size by using several scale, as in the MSR, in order to achieve more robust results, Wang et al. [40].

3.2.5 Discrete cosine transform (DCT)

Chen et al. [5] proposed an alternative technique to normalize illumination. This approach is employed to compensate for illumination variations by truncating the low frequency components of the DCT in the logarithm domain.

3.2.6 Wavelet-based normalization algorithm (WAN)

Du and Ward [8] presented a preprocessing technique based on the wavelet transform. The authors proposed to apply histogram equalization to the sub-band image generated by the so-called approximation wavelet-coefficients and to emphasize the remaining sub-bands generated by the detail-wavelet coefficients. Through the histogram equalization step, the image contrast is improved while the second step enhances the edge information. The final compensated image is obtained by simply employing the inverse wavelet transform on the modified coefficient sub-bands.

3.2.7 Normalization histogram (NH)

Other photometric normalization techniques can also be based on histogram equalization. Histogram equalization is usually achieved by equalizing the histogram of the image pixel gray-levels in the spatial domain so as to redistribute them uniformly. Struc et al. [39] proposed the possibility to replace it with an arbitrary distribution such as the normal, the lognormal and the exponential.

3.2.8 Markov random fields method and models

Markov random field (MRF) is a special type of stochastic process, which was originally used in statistical physics. It is considered as a powerful tool for modeling images and coping with high-dimensional inverse problems from low-level vision, Perez [31]. When MRF is used in the image-processing field, the image is assumed to be segmented of an unknown number of regions, each modelled as individual MRFs. This is achieved by dividing the image into windows. The resulting segmentation (grid) is then used to estimate model parameters. The key concept of the Markov property is the neighborhood, which states that the probability distribution for a particular grid cell value only depends on the cell values within the neighbourhood of the grid cell. The probability model is often formulated in terms of potential functions of the grid values and their choice is a crucial point for the convergence of the method. The use of MRF involves iterative algorithms that in theory converge to the exact probability distribution, however, they might need many iterations so they are often too slow in practical applications and therefore critical for real time applications, Kjønsberg

et al. [18]. Although real-time processing capabilities are very hard to be implemented, they could be achieved by massively parallel hardware architectures, which efficiently exploit the inherent algorithmic parallelism of statistical image models on Markov Random Fields [38].

3.3 Discussion on the photo-normalization techniques for real-time applications

All the photonormalization techniques considered in this work exhibit advantages and drawbacks, especially if used in real time applications. Hereinafter, a brief discussion on pros and cons of each technique is reported. All techniques take advantage of transforming image into the logarithm domain. Since the logarithm of the luminance is a crude approximation of the perceived brightness, logarithm transform can partly reduce the effect of lighting.

3.3.1 Single scale retinex

The Retinex is an image enhancement algorithm that improves the brightness, contrast and sharpness of an image. It performs a non-linear spatial/spectral transform that provides simultaneous dynamic range compression and color constancy. Computation of the Retinex algorithm involves performing a large number of complex operations and data transfers, at video frame rates. In the case of small format images, standard general purpose computers provide sufficient processing power and reasonable performance. When applied to images acquired at real-time video data rates ranging from 15 to 30 fps, a substantial increment in processing speed afforded by hardware performance is required. In addition, several potential applications, as eye tracking, limit the hardware solutions to low-power, low cost, embedded systems, Hines et al. [14]

3.3.2 Multi scale retinex

In the multi-scale version, the equation for single-scale is repeated several times for Gaussians filters with different kernel sizes. The multiscale retinex version is then simply, the weighted sum of a set of single-scale retinex images. The main conceptual problem is that a number of image-processing tasks are performed simultaneously without sufficient regard to the interactions occurring between them. The main practical consequence of this is that MSR is not appropriate for applications which are sensitive to color. A more serious problem is that the implementation of the gray world algorithm is not optimal. In addition, when the algorithm makes use of a large scale implies wide illumination uniformity, but the use of smaller scales yields poor colour constancy results due to local violations of the gray world assumption, and leads to a grayed out image. Averaging the results mitigates the errors, but also reduces the chances for good performance, and thus it is often unsatisfactory. In real-time applications, MSR suffers from the same problems as the SSR and moreover the number of the computational operations increases with the number of the scales used, Barnard and Funt [4].

3.3.3 Single scale selfquotient image

SQI is designed for dealing with the illumination variation especially in face recognition. It requires only one template image for each person because it assumes the same 3D geometry for all persons. In the SQI method the illumination is eliminated by division over a smoothed version of the image itself. It is very simple and can be applied to any single image. However, the weighted Gaussian filter used in SQI can be uneffective in keeping sharp edges in low frequency illumination fields, and the parameter selection is empirical and complicated. For real-time applications, the computational load depends on both the image size and the empirical parameters, as well as on the filter scale, in the sense that it is necessary a specific filter for every convolution window.

3.3.4 Multi scale selfquotient image

The technique exhibits similarities to the MSR technique, but unlike that it uses an anisotropic filter for the smoothing operation. The question of how to scale up these techniques for use with larger numbers of objects still remains unanswered. In real time applications, this technique suffers from the same drawbacks as both SQI and MSR put together.

3.3.5 Discrete cosine transform

The salient feature of DCT is that it does not need any training or modelling step and bootstrap sets and can be easily implemented for high speed computation. There remains another issue: which and how DCT coefficients should be discarded to obtain the well normalized image. Accordingly, illumination variations of images can be reduced by diagonally discarding these low-frequency coefficients. Nevertheless, shadowing and specularity problems are not perfectly solved because they lie in the same frequency band as some features.

3.3.6 Wavelet-based normalization algorithm

Wavelet transform performs multiscale analysis capability, with good edge-preserving ability in low frequency illumination fields. In addition, wavelet decomposition get different band information of the images. It has the advantage of taking into account both contrast and edge enhancements simultaneously.

3.3.7 Normalization histogram

Histogram equalization increases the global contrast of the image while simultaneously compensating for the illumination conditions present at the image acquisition stage, it represents a useful preprocessing step, which can ensure enhanced and more robust recognition performance. While histogram equalization re-maps the histogram of a given image to a uniform distribution, the target distribution could easily be replaced with an arbitrary one. There is no theoretical justification of why the uniform distribution should be preferred to other target distributions. It is a technique simply implementable in real time applications, thanks to the low computational work load.

3.3.8 General discussion on computational load

All the above methods use mainly logarithmic transformation of pixel values, and the convolution of the input image (entirely of a region) with surrounding filters. The filter sizes are empirically chosen to obtain a good tradeoff between real time performance (a small kernel size filter makes the method very fast, Holappa et al. [15]) and pupil tracking accuracy. In general, when small kernel size filters are used the intrinsic information of the image is severely reduced while using large kernel size halo effects might appear. However, single and multiscale retinex (SSR and MSR), SQI and MSQI methods are computationally expensive but a possible good accuracy could be achieved to be used as off-line post processing algorithms. On the contrary, DCT already was used to implement real-time face recognition, Chen et al. [5], could provide good time performance as well as wavelet-based technique. The histogram normalization is a simple equalization, therefore it is a very fast technique. All techniques here suggested are implemented in a modified version of the Matlab INface toolbox.

3.4 Pupil tracking

Pupil tracking algorithm extracts the contour of the pupil exploiting the higher contrast of the pupil than the background due to the IR components of the natural light. Figure 5 shows the algorithm block diagram. More in detail, the first block applies a binarization of the image, obtained by using a threshold, which can be chosen according to two different criteria. Figure 6 reports the histogram of the eye, i.e. the distribution of the image pixel vs the gray levels from 0 to 255. The threshold should divide the histogram into two groups of pixels having only two levels of gray; the zero level should group all the pixel belonging to the pupil whereas the 255 level should identify the background. The first criterion implies choosing the threshold as the absolute minimum value in the range comprised between the two highest peaks of the eye histogram as reported in Fig. 6. The latter criterion implies taking the threshold as the level of gray at which the cumulative sum of pixels, starting from 0, achieves the 3 % of the total pixels of the image. This value was experimentally estimated in order to discriminate the pupil at best. An example of the binarization process is reported in Fig. 7. The group of second, third and fourth block constitutes the preliminary pupil contour detection. More specifically, the second block identifies two sheafs of lines starting from the middle points of the vertical sides of the image, with an angular aperture of 30°. As result of the

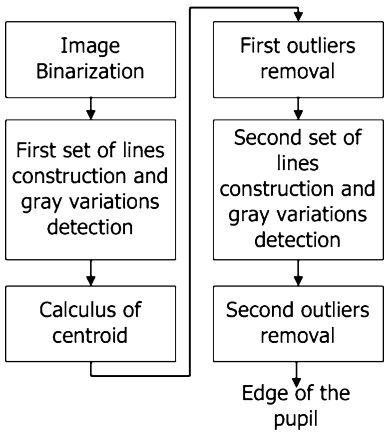


Fig. 5 Block diagram of the pupil tracking algorithm

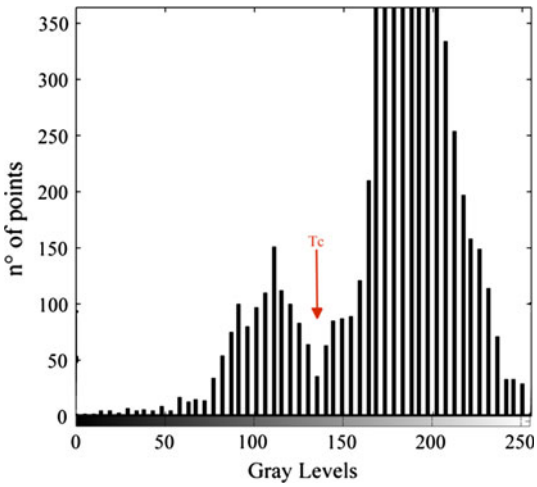
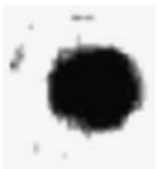


Fig. 6 Example of the histogram of the eye: T_c refers to the threshold identifying the eye and the sclera region

Fig. 7 Example the eye image after the binarization process



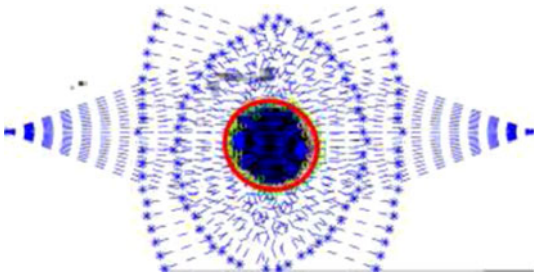


Fig. 8 Application of the pupil tracking algorithm. More specifically, in *black* are the set of points of the eye including the outliers; in *blue* are represented the various lines constructions; in *orange*, the pupil contour are highlighted

binarization process, the image borders are expected to belong to the background, therefore the starting point of each line has a value of 255 in terms of gray level. Analogously, the pupil is expected to be placed roughly in the middle of the image (this is assured by an accurate free-hand selection of the ROI). When each line encounters, along its path, a dark pixel, this latter can be thought to belong to the contour of the pupil. The third block calculates the centroid of these points. The fourth block removes all the outliers, being these points very far from the centroid with respect to the large point density (pupil edge). At this stage, the set of the obtained points constitutes a large-grain approximation of the contour. The fifth block identifies a sheaf of lines starting from the centroid with an angular aperture of 360°, and detects all discontinuities, but now from black to white. The sixth block is another removal operation of the outliers. The result of this algorithm is a set of points constituting the pupil edge. This set will be the input of the fitting algorithm (see Fig. 8).

3.5 Ellipse fitting

Ellipse fitting algorithm is implemented for pupil contour reconstruction, and for detecting the center of the eye. Ellipse is considered as the best geometrical figure representing the eye, being the eye image captured by the camera a projection of the eye in the mirror. According to the ellipse construction, it can be expressed by an implicit second order polynomial, being a central conic (with $b^2 - 4ac < 0$), such as:

$$F(x, y) = ax^2 + bxy + cy^2 + dx + ey + f = 0 \quad (8)$$

Ellipse fitting algorithms present in the literature can be divided into two broad techniques: the clustering/vot-ing (CV) and the least square (LS) techniques. The first one uses two main approaches, such as RANSAC and Hough Transform. RANSAC technique is extremely robust but it is time-demanding and memory-consuming, Forsyth and Ponce [11], and being an iterative algorithm does not have fixed computation time therefore, it is not suitable for real time applications. The Hough Transform suffers from sensitivity limitations due to the presence of spurious and blurred peaks, Grimson and Huttenlocher [13], and even though many efforts were made for computational cost reduction, the algorithm seems still to be excessively resource consuming for real time machine vision, Bennett et al. [3]. The LS method is based on finding a set of parameters that minimize the distance between the data points and the ellipse. According to literature, this technique fulfills the real time issue. One implementation of LS technique has been introduced by Fitzgibbon et al., which is a direct computational method (i.e. B2AC, the acronym is based on the well-known discriminant formula for quadratic polynomials) based on the algebraic distance with a quadratic constraint, Fitzgibbon et al. [10]. In this work, we use a custom B2AC algorithm, where a gaussian noise is added for algorithm stabilization, Maini [25], to calculate the center of the pupil (that coincides with ellipse center), the axes dimensions as well as the eccentricity.

3.6 Mapping of the position of the eye

The mapping procedure associates the eye center position to the image plane of the scene, providing as result the gaze point, see (see Fig. 9). An experimenter guides this procedure. Firstly, the camera is positioned to capture both the scene (in our case the screen) and the mirror. In detail, tilt of the camera is adjusted as well as the tilt of mirror shaft and the tilt of the mirror to reflect the eyes. Each participant is asked to look at some specific points of the screen. These points are identified by coordinates $s_i = (x_{si}, y_{si})$ referred to the image plane (i.e. the image plane captured by the camera), (see Fig. 3). The participants were instructed to keep their head as still as possible and to carefully look at each target point without blinking until looking at the next one.

The mapping function gets as input the center of the eye coming from Ellipse fitting block, and the coordinates of point on the image plane.

Mapping functions are quadratic polynomials defined as:

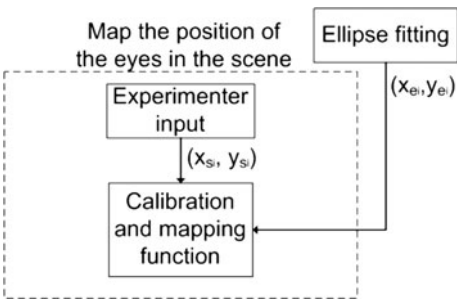


Fig. 9 Block Diagram of the mapping function calculation process

$$x_{si} = a_{11} + a_{12}x_{ei} + a_{13}y_{ei} + a_{14}x_{ei}y_{ei} + a_{15}{x_{ei}}^2 + a_{16}{y_{ei}}^2 \quad (9)$$

$$y_{si} = a_{21} + a_{22}x_{ei} + a_{23}y_{ei} + a_{24}x_{ei}y_{ei} + a_{25}{x_{ei}}^2 + a_{26}{y_{ei}}^2 \quad (10)$$

where x_{si}, y_{si} are the coordinates of the image plane (i.e. the coordinates of the point on the screen mapped into the image plane captured by the camera), and x_{ei}, y_{ei} are the coordinates of the center of the eye coming from the ellipse fitting block, referred to the image plane as well. The coefficients $a_{1,1-6}$, and $a_{2,1-6}$ are unknown. Since each calibration point defines two equations, the system is over constrained with 12 unknowns and 18 equations, and can be solved using least square method (LSM).

3.7 Experimental setup

Ten subjects, nine males and one female, volunteered to participate in the experiment. Six subjects had dark eyes and four had bright eyes. The average age was of 26.8 with a standard deviation of 1.5. The experiment was performed in three illumination conditions: darkness, laboratory and sunlight condition. In detail, the laboratory condition was achieved by white neon lighting equally distributed in the room with a power of 50 lumens. Dark condition is achieved by reducing by 55% the power of the illumination, and the sunlight condition is obtained performing the experiment at 12:30 a.m. at the top floor of the university building in a room having two walls completely windowed. The experimental protocol was structured in three phases. In the first phase the participant was instructed on the experiment modalities. In the second phase calibration procedure was performed, and finally the experimental test was carried out. The whole experiment lasted about 8 min; the first 4 min were used to describe the experiment and the remaining 4 for calibration and testing part. More in detail, during the first phase all the subjects were asked to sit on a comfortable chair in front of a screen at a fixed distance of 70 cm, while they wear the HATCAM system. The system was also equipped with a chin-support in order to avoid head movement. In the calibration procedure, the system was tailored (position and orientation of mirror and camera) to the user characteristic profile. During the experimental test, all subjects were asked to look at the point which appears on the screen, kept, one by one, for 2.5 s. Each subject was presented with 36 randomly distributed points for each illumination condition (see Fig. 10), for a grand total of 108 points. The sequence of illumination conditions is randomly changed for each subject to avoid systematic errors.

The different processing methods were tested in Matlab environment running on a hardware platform based on 2.2 GHz Dual Core with 2G memory computer. In order to estimate run time per frame, we used a dedicated function interrogating the timer.

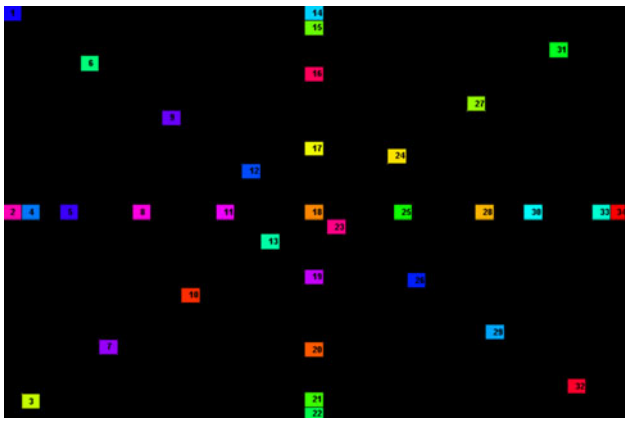


Fig. 10 Example of the points showed on the screen for accuracy EGT system evaluation

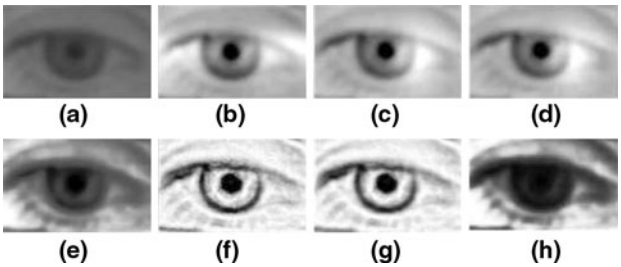


Fig. 11 Example of all photonormalization techniques applied to the same gray-eye image. In detail: **a** original image, **b** reconstructed image by applying DCT, **c** Reconstructed image by applying SSR, **d** Reconstructed image by applying MSR, **e** reconstructed image by applying NH, **f** reconstructed image by applying SQI, **g** reconstructed image by applying MSQ, **h** reconstructed image by applying WAN

3.8 Results and discussion

The aim of this work was to compare the performance of seven photonormalization algorithms on pupil area detection and gaze tracking under different illumination conditions. Results are reported and discussed in terms of errors between the estimated and real measurements. The performance evaluation of the different algorithms took into account the execution time as well. More specifically, Fig. 11 shows an example of the all photonormalization techniques applied to the same gray-level image of an eye. Although the algorithms provide results whose differences can be easily detected at a glance, a more quantitative comparison has to be done. We performed a careful statistical analysis of errors calculated as difference between real and estimated measurements for eye tracking and pupil area detection. The errors, in terms of pixels, computed

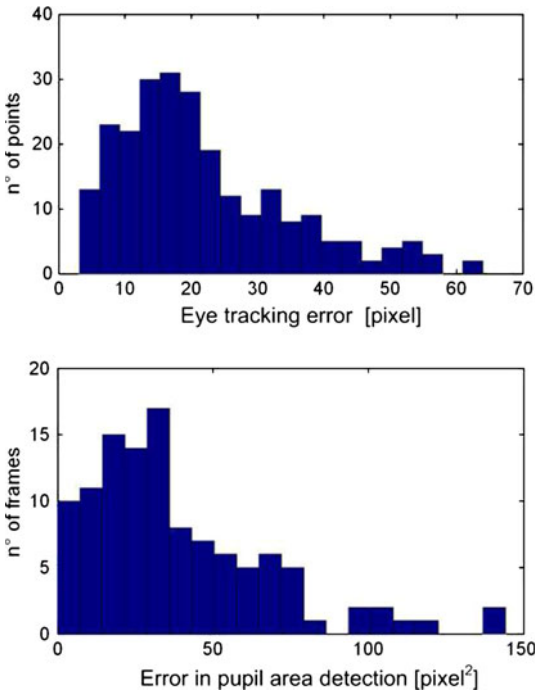


Fig. 12 Example of the distributions of errors in eye tracking and pupil area detection

through all the subjects and frames, are not normally distributed. In Fig. 12, by way of example, the error histograms, for one algorithm and illumination condition, are reported.This is also confirmed by the Lilliefors [24] test, which returns a p value rejecting the null hypothesis of normality in all the cases. Accordingly, we used the Kruskal and Wallis [19] (Kruskal–Wallis) test, which is a non parametric one-way analysis of variance by ranks for testing equality of population medians. Like most non-parametric tests, Kruskal–Wallis is performed on ranked data, so the measurement observations are converted to their ranks in the overall data set. This test does assume an identically shaped and scaled distribution for each group, except for any difference in medians. The null hypothesis is stated as the probability that the samples come from identical populations, regardless their distributions.

In place of the mean of distributions, we considered the median as a measure of location, usually taken when the distribution is skewed, such as in our case (see Fig. 12), Stavig and Gibbons [37]. Tables 1 and 2 show median and dispersion of each photonormalization technique in the three illumination conditions for both eye tracking and pupil area detection. The lowest values of median are reported in bold. As it is well-known from the statistical analysis, in order to evaluate performance of the seven photonormalization techniques, it is not sufficient to compare the means rather than medians (not being normal distributions) of the groups, but we have to assess whether the groups statistically come from different populations. Applying the Kruskal–Wallis test to all the possible combinations among the seven techniques relatively to eye tracking and for each illumination condition we obtained four statistically equivalent classes of techniques, i.e. the null hypothesis in each class cannot be rejected. In Table 3, four different classes are identified for each illumination condition. The classes are sorted out by increasing median, therefore the best photonormalization techniques belong to the class 1, as it exhibits the lowest median of error distributions. The techniques in common with all the three illumination conditions are the NH and DCT. Similarly, we applied the same methodology for the pupil area error distributions and results are reported in Table 4. In this case, we identified three classes for each illumination condition. Following the same approach as above, the technique in common with the three illumination conditions is the DCT. In addition to the error statistical analysis, we also considered the execution time of the seven algorithms under the three different illumination conditions. Even in this case, the Lilliefors test was applied to the distributions of the execution time calculated for each frame of the video recordings, for all the photonormalization techniques and illumination conditions, giving as result that all the distributions are not normal. In order to compare the execution time, median and dispersion of execution time distributions for all the seven techniques were calculated and reported in the Table 5. It is worthwhile noting that the MSQ technique takes much more time than the other ones, and it is not suitable for real time applications. This is due to the fact that, in order to achieve an accuracy in pupil detection and gaze tracking comparable to the other techniques, it requires a larger kernel size filter with a subsequent higher computational load. These values reported in the table do not change under the three illumination conditions. The fastest algorithm resulted to be the NH. The choice of the most suitable technique depends on the need of having an accurate rather than fast

Table 1 Median and dispersion of eye tracking errors in pixels

	SSR	MSR	SQI	MSQ	DCT	WAN	NH
Laboratory	19.686 ± 10.771	22.091 ± 12.243	28.234 ± 13.339	28.108 ± 13.979	16.214 ± 7.016	23.853 ± 8.221	17.114 ± 7.364
Sunlight	18.712 ± 7.045	18.067 ± 7.384	32.375 ± 17.215	34.064 ± 17.731	16.522 ± 6.641	26.036 ± 9.135	16.053 ± 6.235
Darkness	23.037 ± 12.908	21.415 ± 11.870	29.052 ± 13.874	34.368 ± 19.000	16.051 ± 6.595	22.504 ± 5.64	17.599 ± 7.229

Table 2 Median and dispersion of pupillometry errors in $pixel^2$

	SSR	MSR	SQI	MSQ	DCT	WAN	NH
Laboratory	35.847 ± 24.253	36.024 ± 21.287	36.231 ± 26.491	35.688 ± 23.711	30.815 ± 20.722	47.547 ± 26.264	39.092 ± 15.661
Sunlight	48.884 ± 23.126	49.377 ± 21.06	25.381 ± 17.999	30.643 ± 17.589	35.716 ± 18.172	33.489 ± 21.401	76.348 ± 20.696
Darkness	45.971 ± 24.770	46.981 ± 26.079	69.495 ± 33.799	71.424 ± 36.997	41.632 ± 26.202	110.695 ± 34.020	37.557 ± 16.675

Table 3 Eye tracking error classes

	Class 1	Class 2	Class 3	Class 4
Laboratory				
Techniques	DCT, NH	SSR	MSR	SQI, MSQ, WAN
p value	0.6242			0.3399
Sunlight				
Techniques	DCT, NH	SSR, MSR	WAN	SQI, MSQ
p value	0.5921	0.7577		0.9077
Darkness				
Techniques	DCT, NH	SSR, MSR, WAN	SQI	MSQ
p value	0.3426	0.6688		

Table 4 Pupillometry error classes

	Class 1	Class 2	Class 3
Laboratory			
Techniques	DCT, SSR, MSR	NH, SQI, MSQ	WAN
p value	0.3548	0.9478	
Sunlight			
Techniques	DCT, SQI, MSQ, WAN	SSR, MSR	NH
p value	0.9512	0.6143	
Darkness			
Techniques	DCT, SSR, MSR, NH	SQI, MSQ	WAN
p value	0.1364	0.9400	

Table 5 Median and dispersion of the execution time (s) intended as per-frame processing time

SSR	MSR	SQI	MSQ	DCT	WAN	NH
0.0175 ± 0.0011	0.0407 ± 0.0018	0.3028 ± 0.0199	1.3700 ± 0.0496	0.0224 ± 0.0019	0.0255 ± 0.0050	0.0098 ± 0.0008

algorithm. From Table 3 we can state that, for eye tracking in all the illumination conditions, DCT and NH are the best photonormalization techniques. Reading the first column of Table 4 we derive that for the pupil area detection in laboratory condition DCT, SSR and MSR are the best techniques; in sunlight condition DCT, SSQ, MSQ and WAN show the best results and in darkness condition DCT, SSR, MSR and NH exhibit the lowest error. These results show that DCT is one of the best techniques in pupil area detection, and it is also very effective for eye tracking. NH shows good results for eye tracking in all conditions and also in darkness condition for pupillometry. The remaining techniques are, in general, less effective under the three illumination conditions both in eye tracking and pupillometry. If we take into account also the execution time the best technique resulted to be NH, which is executed in less than 1 ms followed by SSR which is executed in less than 18 ms (but effective only for pupillometry during laboratory and darkness conditions) and third is DCT which is executed in less than 23 ms. The remaining techniques are slower.

As DCT resulted to be the best technique, it was optimized in C++ language using *openframeworks* library. In this case, the execution time per frame resulted to be much lower than that calculated in Matlab environment as reported in Table 5. In addition, the execution time of the

Table 6 Execution time of the different stages of a whole algorithm based on DCT photonormalization technique

Task	Time (ms)
DCT	3.25
Beam tracking	21.26
Fitting ellipse	1.25
Total software	34.47

whole algorithm for eye tracking was computed in detail and reported in Table 6 using the *openframeworks* command *ofGetElapsedTimeMillis*() which gives the elapsed time expressed in milliseconds. It is worthwhile pointing out that the reported execution time are compatible for real time applications and future works will be addressed to implement this algorithm in embedded systems based on linux OS such as *FOXBOARD*.

3.9 Conclusions

In this paper, we investigated on seven photonormalization techniques in order to increase the robustness of the HATCAM eye tracker. Our key approach was to evaluate, in terms of accuracy of eye tracking, pupillometry and execution time, seven photonormalization techniques under three different brightness conditions: laboratory, sunlight, darkness. The results showed that the DCT method is the most effective in terms of accuracy of gaze point and pupillometry. The NH is the best in terms of eye tracking but is less effective in the pupillometry; nevertheless it results to be the fastest. The execution time is a crucial issue for the real time processing. Moreover, a fast processing together with a fast camera could allow the system to also detect fast saccadic movements. In conclusion to obtain a good effectiveness in all the illumination conditions this study suggests of using DCT, although when the HATCAM is used only as eye tracker, the preference goes to the NH technique.

References

1. Anon Asl.: Also available as http://www.asl.com (2010)
2. Babcock, J., Pelz, J., Peak, J.: The wearable eyetracker: a tool for the study of high-level visual tasks. In: Proceedings of the Military Sensing Symposia Specialty Group on Camouflage, Concealment, and Deception, Tucson, Arizona, Citeseer (2004)
3. Bennett, N., Burridge, R., Saito, N.: A method to detect and characterize ellipses using the Hough transform. IEEE Trans. Pattern Anal. Mach. Intell. **21**(7), 652–657 (2002)
4. Barnard, K., Funt, B.: Investigations into multi-scale Retinex. In: Colour Imaging Vision and Technology, CIteseer. Wiley, New York, pp. 9–17 (1999)
5. Chen, W., Er, M., Wu, S.: Illumination compensation and normalization for robust face recognition using discrete cosine transform in logarithm domain. IEEE Trans. Syst. Man Cybernet. Part B: Cybernet. **36**(2), 458–466 (2006)
6. DiScenna, A., Das, V., Zivotofsky, A., Seidman, S., Leigh, R.: Evaluation of a video tracking device for measurement of horizontal and vertical eye rotations during locomotion. J. Neurosci. Methods **58**(1–2), 89–94 (1995)
7. Dongheng, L., Winfield, D., Parkhurst, D.: Starburst: a hybrid algorithm for video-based eye tracking combining feature-based and model-based approaches. In: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (2005)
8. Du, S., Ward, R.: Wavelet-based illumination normalization for face recognition. In: IEEE International Conference on Image Processing, 2005. ICIP 2005. IEEE, vol. 2 (2005)
9. Duchowski, AT.: Eye Tracking Methodology. Springer, Berlin (2003)
10. Fitzgibbon, A., Pilu, M., Fisher, R.: Direct least square fitting of ellipses. IEEE Trans. Pattern Anal. Mach. Intell. **21**(5), 476–480 (2002)
11. Forsyth, D., Ponce, J.: Computer vision: a modern approach. Prentice Hall Professional Technical Reference (2002)
12. Franchak, J., Kretch, K., Soska, K., Babcock, J., Adolph, K.: Head mounted eye tracking of infants' natural interactions: a new method. In: Proceedings of the 2010 Symposium on Eye Tracking Research Applications, Citeseer, pp. 21–27 (2010)
13. Grimson, W., Huttenlocher, D.: On the sensitivity of the Hough transform for object recognition. IEEE Trans. Pattern Anal. Mach. Intell. **12**(3), 255–274 (2002)
14. Hines, G.D., Rahman, Z., Jobson, D.J., Woodell, G.A.: Single-scale retinex using digital signal processors. In: Global Signal Processing Conference, CIteseer, 27(30) (2004)
15. Holappa, J., Ahonen, T., Pietikainen, M.: An optimized illumination normalization method for face recognition. In: 2nd IEEE International Conference on Biometrics: Theory, Applications and Systems, 2008. BTAS 2008. pp. 1–6 (2008)
16. Jobson, D., Rahman, Z., Woodell, G.: A multiscale retinex for bridging the gap between color images and the human observation of scenes. IEEE Trans. Image Process. **6**(7), 965–976 (2002)
17. Jobson, D., Rahman, Z., Woodell, G.: Properties and performance of a center/surround retinex. IEEE Trans. Image Process. **6**(3), 451–462 (2002)
18. Kjonsberg, H., Kolbjornsen, O.: Markov mesh simulations with data conditioning through indicator kriging. In: Proceedings of Geostats, vol. 8 (2008)
19. Kruskal, W., Wallis, W.: Use of ranks in one-criterion variance analysis. J. Am. Stat. Assoc. **47**(260), 583–621 (1952)
20. Land, E., McCann, J.: Lightness and retinex theory. J. Optical Soc. Am. **61**(1), 1–11 (1971)
21. Land, M., Lee, D.: Where we look when we steer. Nature **369**(6483), 742–744 (1994)
22. LcTechnologies (2010) Also available as http://www.eyegaze.com
23. Li, D., Babcock, J., Parkhurst, D.: OpenEyes: a low-cost head-mounted eye-tracking solution. In: Proceedings of the 2006 Symposium on Eye Tracking Research and Applications. ACM, pp 95–100 (2006)
24. Lilliefors, H.: On the Kolmogorov–Smirnov test for normality with mean and variance unknown. J. Am. Stat. Assoc. **62**(318), 399–402 (1967)
25. Maini, E.: Robust ellipse-specific fitting for real-time machine vision. In: Brain, Vision, and Artificial Intelligence, vol. 3704, pp. 318–327 (2005)
26. Marsi, S., Saponara, S.: Integrated video motion estimator with retinex-like pre-processing for robust motion analysis in automotive scenarios: algorithmic and real-time architecture design. J. Real-Time Image Process. **5**(4), 275–289 (2010)

27. Meyer, A.B., Ohme, M., Martinetz, T., Barth, E.: A single-camera remote eye tracker. Percept. Interact. Technol. 208–211 (2006)
28. Morimoto, C., Mimica, M.: Eye gaze tracking techniques for interactive applications. Comput. Vis. Image Underst. **98**(1), 4–24 (2005)
29. Morimoto, C., Amir, A., Flickner, M.: Detecting eye position and gaze from a single camera and 2 light sources. Pattern Recogn. **4**(40), 314 (2002)
30. Pelz, J., Canosa, R., Babcock, J.: Extended tasks elicit complex eye movement patterns. In: Proceedings of the 2000 symposium on Eye tracking research & applications, ACM, p. 43 (2000)
31. Perez, P.: Markov random fields and images. CWI Quarterly, Institut de recherche en informatique et syst {è} mes al {é} atoires. **11**(4), 413–437 (1998)
32. Ranawade, S.: Face recognition and verification using artificial neural network. Int. J. Comput. Appl. IJCA **1**(14), 23–30 (2010)
33. Robinson, D.: A method of measuring eye movemnent using a scieral search coil in a magnetic field. IEEE Trans. Biomed. Electron. **10**(4), 137–145 (2008)
34. San Agustin, J., Skovsgaard, H., Mollenbach, E., Barret, M., Tall, M., Hansen, D.W., Hansen, J.P.: Evaluation of a low-cost open-source gaze tracker. In: Proceedings of the 2010 symposium on eye-tracking research & applications, ACM, pp. 77–80 (2010)
35. Shashua, A., Riklin-Raviv, T.: The quotient image: Class-based re-rendering and recognition with varying illuminations. IEEE Trans. Pattern Anal. Mach. Intell. **23**(2), 129–139 (2002)
36. Sliney, D., Aron-Rosa, D., DeLori, F., Fankhauser, F., Landry, R., Mainster, M., Marshall, J., Rassow, B., Stuck, B., Trokel, S. et al.: Adjustment of guidelines for exposure of the eye to optical radiation from ocular instruments: statement from a task group of the International Commission on Non-Ionizing Radiation Protection (ICNIRP). Appl. Optics **44**(11), 2162–2176 (2005)
37. Stavig, G., Gibbons, J.: Comparing the mean and the median as measures of centrality. International Statistical Review/Revue Internationale de Statistique **45**(1), 63–70 (1977)
38. Stilkerich, S., Reiger, R.: On the simulation and development of massive parallel digital architectures for Markov random fields [image processing applications]. In: Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 5, pp.169–172 (2004)
39. Struc, V., Zibert, J., Pavesic, N.: Histogram remapping as a preprocessing step for robust face recognition. WSEAS Trans. Inf. Sci. Appl. **6**(3), 520–529 (2009)
40. Wang, H., Li, S., Wang, Y.: Face recognition under varying lighting conditions using self quotient image. In: Proceedings of Sixth IEEE International Conference on Automatic Face and Gesture Recognition, 2004. IEEE, pp. 819–824 (2004a)
41. Wang, H., Li, S., Wang, Y.: Generalized quotient image. IEEE Computer Society (2004b)
42. Young, L., Sheena, D.: Survey of eye movement recording methods. Behav. Res. Methods Instrum. **7**, 397–429 (1975)
43. Zhu, Z., Ji, Q.: Novel eye gaze tracking techniques under natural head movement. IEEE Trans. Biomed. Eng. **54**(12), 2246–2260 (2007)

Author Biographies

Antonino Armato received the Laurea degree in Electronic Engineering with a specialization in "Systems and Automation" from University of Messina, Italy, on May [25], with degree thesis "Analysis, Modelling and Simulation of the shoulder complex kinematics". From May to October 2005, he worked on the sensors and electronic interface of HEaLthcare Service Linking "TelerehabilitatiOn" to Disable peOple and Clinicias (HELLODOC) project at Signo Motus srl, Messina, Italy. In June 2006, he received the Masters in Information and Communication of Technology at the ICT Center of Excellence For Research, Innovation, Education and Industrial Labs partnership (CEFRIEL)—Polytechnic of Milan, Italy. From September to December 2006, he worked at the Sintesi S.C.p.A., Bari, Italy, about the user interfaces and image processing of Small Animal Tomographic Imaging System (SATIS). From January to December 2007, he was assistant research at the ARTS-Labs Scuola Superiore Sant'Anna. Since January 2008 he is PhD Student in Automatic, Robotic and Bioengineering at the Interdepartmental Research Center "E. Piaggio" of the Faculty of Engineering of University of Pisa. His research interests are in reconfigurable computing, neural networks and eye tracking systems.

Antonio Lanatà Ph.D. graduated in Electronic Engineering from the University of Pisa in 2001, with a dissertation on electro-active polymers for actuator systems. In December 2006, he received the Ph.D degree in Automation, Robotics and Bioengineering at University of Pisa (Italy), with a dissertation on the innovative wearable systems for heart monitoring. In 2007, he received a postdoctoral fellowship on innovative cardiopulmonary systems monitoring. His research interests are focused on wireless wearable systems for health status monitoring by ultrasound-acoustic, UltraWideBand technologies and signal processing.

Enzo Pasquale Scilingo Ph.D. is Assistant Professor in Electronic and Information Bioengineering at the University of Pisa. He received the Laurea Degree in Electronic Engineering from the University of Pisa, Italy and the Ph.D. degree in Bioengineering from the University of Milan, in 1995 and 1998 respectively. For 2 years, he was post-doctoral fellow with the Italian National Research Council and for 2 years post-doctoral fellow at Information Engineering Department of the University of Pisa. Currently, he is pursuing his research work mainly at the Interdepartmental Research Center E. Piaggio. His research interests are in biomedical and biomechanical signal processing, modelling and control and instrumentation. He is author of several papers on peer-review journals, contributions to international conferences and chapters in international books.